Outcomes reported by the Vascular Quality Initiative and the National Surgical Quality Improvement Program are not comparable

Francesco A. Aiello, MD, Bing Shue, MD, Nisha Kini, MBBS, MPH, Amy Rosen, PhD, Louis Messina, MD, William Robinson, MD, Philimon Gona, PhD, and **Andres Schanzer, MD,** *Worcester, Mass*

Objective: **The Vascular Quality Initiative (VQI) and National Surgical Quality Improvement Program (NSQIP) have emerged as the primary vascular surgery quality measurement tools with the purpose of evaluating perioperative outcomes and assessing hospital and physician quality. VQI uses self-reporting to capture all index vascular procedures during the inpatient period. NSQIP employs nurse abstractors to capture a sample of procedures and covers 30-day events. We hypothesize that patients undergoing lower extremity bypass (LEB) will exhibit high concordance for preoperative variables and low concordance for postoperative variables between these data sets.**

Methods: **All patients undergoing LEB for peripheral arterial disease at the University of Massachusetts captured in both VQI and NSQIP databases were reviewed (2007-2012). Concordance between categorical variables was assessed by κ correlation coefficient. All postoperative variables were compared during equivalent inpatient stay. Events between discharge and 30 days postoperatively were tabulated with use of the NSQIP data set.**

Results: **We identified 240 patients undergoing LEB captured in both VQI and NSQIP. Comparison of this identical patient cohort between VQI and NSQIP revealed a moderate to strong agreement for most preoperative variables except for congestive heart failure ($\kappa = 0.14$) and hypertension ($\kappa = 0.35$), which showed poor agreement. Concordance for inpatient postoperative variables was high for mortality ($\kappa = 1.0$) and myocardial infarction ($\kappa = 0.86$) but moderate for pulmonary complications ($\kappa = 0.57$) and poor for unplanned return to the operating room ($\kappa = 0.41$), wound infection ($\kappa = -0.01$), and change in renal function ($\kappa = -0.01$). A majority of postoperative events (71%) occurred between discharge and 30 days postoperatively, with a significantly higher incidence of wound infections in the outpatient setting (4.2% vs 95.8%; $P < .0001$).**

Conclusions: **VQI and NSQIP demonstrate substantial concordance for most preoperative variables and poor concordance for most postoperative variables, even at identical collection periods. This discordance is a result of differences in data collection methods and variable definitions. On the basis of these findings, VQI and NSQIP data sets cannot be used to directly compare risk-adjusted patient outcomes between institutions. (J Vasc Surg 2014;60:152-9.)**

The U.S. health care landscape continues to shift as the medical community braces itself for the application of the Patient Protection and Affordable Care Act (ACA) of 2010 and subsequent potential changes in the Medicare reimbursement program.[1] The proposed penalties and net effect of the ACA reduction programs for hospital readmissions and hospital-acquired conditions remain uncertain, but with the proposed inclusion of "vascular conditions" in fiscal year 2015, the effect on physicians performing vascular surgery procedures could be significant.[2,3] Accurate and detailed reporting of patient demographics, comorbidities, procedural details, and outcomes will be crucial to avoid unintended penalties and reductions in Medicare reimbursement.[4]

Risk adjustment based on specific patient factors is not novel to outcomes reporting in vascular surgery, especially in the area of lower extremity revascularization for peripheral arterial disease.[5,6] Multiple large studies leveraging data collected from multiple centers have attempted to derive and to validate risk-adjusted prediction models to predict postoperative outcomes.[5-9] Whereas these studies have provided prediction models for long-term outcomes including amputation-free survival and mortality, these efforts have not addressed immediate perioperative outcomes. Databases have only recently been employed to measure perioperative surgical quality and will continue to play an increasingly important role.[10]

In the United States, two independently maintained quality improvement databases have emerged as the primary vascular surgery quality measurement tools with the purpose of evaluating perioperative outcomes and assessing hospital and physician quality. These two databases are the Vascular Quality Initiative (VQI) and the American College of Surgeons National Surgical Quality

From the Division of Vascular and Endovascular Surgery, University of Massachusetts Medical School.

Author conflict of interest: none.

Presented at the plenary session of the Fortieth Annual Meeting of the New England Society for Vascular Surgery, Stowe, Vt, September 27-29, 2013.

Additional material for this article may be found online at www.jvascsurg.org.

Reprint requests: Francesco A. Aiello, MD, Assistant Professor of Surgery, Division of Vascular and Endovascular Surgery, University of Massachusetts Medical School, 55 Lake Ave N, S3-830, Worcester, MA 01655 (e-mail: francesco.aiello@umassmemorial.org).

The editors and reviewers of this article have no relevant financial relationships to disclose per the JVS policy that requires reviewers to decline review of any manuscript for which they may have a conflict of interest.

0741-5214/$36.00

http://dx.doi.org/10.1016/j.jvs.2014.01.046

Improvement Program (NSQIP). Although derivation and validation of risk prediction models have been performed separately in both VQI and NSQIP, a comparison of the variables and outcomes as captured and reported in each database has not been performed to date.[9,10] This could have a dramatic impact on results obtained across institutions and the ability to benchmark postoperative outcomes if different data sets are used for comparison.

The purpose of this study was to compare the capture of clinical data for an identical patient cohort undergoing lower extremity bypass (LEB) and to quantify the concordance for preoperative and postoperative variables, as recorded by each independent database. Preoperative variables were obtained on the basis of chart reviews and existing medical records, which tend to be more robust in established patients with multiple clinical visits. This is in stark contrast to postoperative events, where such an episode must be new and occur within a given postoperative time frame, lending itself to potential collection bias and all the challenges that are associated with postoperative data collection. We therefore hypothesized that preoperative variables will be highly concordant between databases and that postoperative variables will be discordant between databases.

METHODS

Study design. This study was a retrospective comparison of preoperative and postoperative variables as recorded by the VQI and NSQIP databases in patients who underwent an infrainguinal LEB at the University of Massachusetts Medical Center. All patients who underwent an infrainguinal LEB between 2007 and 2012 and who were captured by both VQI and NSQIP databases were included in the analysis. All patients were matched across the two databases by the unique medical record number, date, and type of procedure to ensure that the same individual patient's data were compared in the two databases. Any patient not matched by each of these measures was excluded from the study. Multiple bypass procedures on a single patient were counted separately if the second procedure did not include any portion of the previous bypass.

Concordance between preoperative and postoperative variables between the two databases was assessed. To ensure that all comparisons were based on a uniform data collection time frame, postoperative events were compared between VQI and NSQIP for the inpatient period only. As a secondary subset analysis for patients in the NSQIP data set, we compared the proportion of postoperative events that occurred during the inpatient period to those that occurred in the period after discharge through 30 days after surgery. The University of Massachusetts Medical Center Institutional Review Board approved this study protocol.

Databases. The VQI is a national cooperative quality improvement initiative incorporating both academic and community hospitals designed to evaluate processes of care and outcomes in vascular surgery.[11] All data are self-reported at each participating institution by physicians, nurses, or clinical personnel with use of standardized definitions. Patient demographics, comorbidities, and intraoperative and postoperative information are collected for the inpatient period from the index case. Additional data collected at 1 year after the procedure are included. VQI has 100% capture for 11 major vascular surgery procedures. Aggregate data collection is used to generate benchmark reports to assess quality of care with outcomes data; the VQI program currently includes 242 participating centers.

NSQIP is a national database developed by the American College of Surgeons to collect preoperative through 30-day postoperative data. Surgical cases are sampled on the basis of institution program options with a minimal requirement of cases analyzed annually. The essential program, as selected for this study, involved cases selected every 8 days to ensure appropriate sampling.[12] Data are collected on the basis of chart reviews by surgical clinical reviewers trained by the American College of Surgeons. Surgical clinical reviewers are responsible for collecting patient information at each institution based on preset standardized definitions with routine audits performed to ensure accurate collection. Risk-adjusted information is provided to all participating hospitals to help benchmark complication rates and surgical outcomes.

Variables. Preoperative demographic and clinical variables included gender, race, ethnicity, presence of diabetes, current smoking, hypertension, congestive heart failure (CHF), dialysis dependence, chronic obstructive pulmonary disease (COPD), and American Society of Anesthesiologists (ASA) physical status classification.

Postoperative outcomes included mortality, myocardial infarction (MI), change in renal function, pulmonary complications (pneumonia and unplanned reintubation), return to the operating room, and wound infection (Appendix, online only).

Definitions. Most variables were defined similarly between the two databases. There were no differences in the definitions of sex, race, ethnicity, and dialysis dependence between data sets. Similarly, in both data sets, the variable diabetes was stratified as nondiabetic, non–insulin-dependent diabetes, or insulin-dependent diabetes; the variable smoking was defined as nonsmoker (no tobacco use within the prior 12 months) and current smoker (tobacco use within the previous 12 months). All patients with a diagnosis of COPD, CHF, and hypertension were considered positive for the disease if it was documented in either data set regardless of severity. The ASA score was included on the basis of criteria determined by the ASA for preoperative risk assessment as recorded in either data set (see Appendix, online only, for full definition).

For any variables with different definitions between VQI and NSQIP, a single standardized definition was created to allow a fair comparison across the two databases. For example, VQI captures four different levels of CHF, whereas NSQIP has a yes or no answer based on a single inclusion definition. In this case, all levels of CHF from

Table I. Comparison of preoperative variables in National Surgical Quality Improvement Program (*NSQIP*) and Vascular Quality Initiative (*VQI*) data sets for infrainguinal bypass (total N = 240)

	NSQIP		*VQI*		*κ statistic*	
Variable	*No.*	*%*	*No.*	*%*	*κ coefficient*	*P value*
Gender						
Male	157	65.4	159	66.3	0.96	<.0001
Female	83	34.6	81	33.8		
Ethnicity						
Hispanic or Latino	10	4.3	7	3.0	0.82	<.0001
Non-Hispanic or Latino	225	95.7	228	97.0		
Race						
White	219	91.6	228	95.4	0.55	<.0001
Nonwhite	20	8.4	11	4.6		
Diabetes mellitus						
No	135	56.3	125	52.1	0.74	<.0001
Oral medication/noninsulin	39	16.3	57	23.8		
Insulin	66	27.5	58	24.2		
Smoking						
Yes	78	32.5	86	35.8	0.69	<.0001
No	162	67.5	154	64.2		
CHF						
Yes	4	1.7	32	13.3	0.14	.0003
No	236	98.3	208	86.7		
COPD						
Yes	35	14.6	33	13.8	0.49	<.0001
No	205	85.4	207	86.3		
Hypertension						
Yes	204	85.0	197	82.1	0.35	<.0001
No	36	15.0	43	17.9		
Dialysis						
Yes	6	2.5	7	2.9	0.76	<.0001
No	234	97.5	233	97.1		
ASA class						
1	0	0.0	0	0.0	0.58	<.0001
2	15	7.3	25	12.1		
3	177	85.9	170	82.5		
4	14	6.8	11	5.3		
5	0	0.0	0	0.0		

ASA, American Society of Anesthesiologists; *CHF,* congestive heart failure; *COPD,* chronic obstructive pulmonary disease.

VQI were categorized as yes to be compared with the CHF variable as recorded by NSQIP (see Appendix, online only, for full definitions).

Postoperative variables were considered categorical yes/no variables and not further stratified according to the severity of the condition. All variables were documented according to the exact way in which they were recorded in the data set.

Statistical analysis. Concordance between all categorical variables collected in VQI and NSQIP was assessed by the κ correlation coefficient. The κ correlation coefficient measures inter-rater agreement or inter-annotator agreement for categorical items.[13] A high κ (>0.75) indicates strong agreement, and a low κ (≤0.40) indicates poor agreement.[a] The Spearman rank correlation coefficient, *r*, was used to assess agreement between all ordinal variables. Spearman rank correlation coefficient measures correlation between ranked variables. Higher Spearman rank correlation coefficient indicates higher correlation between ranked variables. A *P* value < .05 was considered statistically significant. All analyses were conducted with SAS 9.3 (Cary, NC) and STATA 12.1 (College Station, Tex).

RESULTS

VQI identified 564 LEB patients, of whom 324 were excluded because they were not captured by NSQIP. NSQIP identified 299 patients; 54 of these patients were excluded because they were not also present in VQI. Overall, 240 patients found in both the VQI and NSQIP data sets were analyzed.

Comparison of preoperative variables between VQI and NSQIP. Preoperative characteristics of patients undergoing infrainguinal LEB were compared between the VQI and NSQIP data sets (Table I). Patient gender (κ = 0.96) and ethnicity (κ = 0.82) were in almost perfect

[a]κ correlation coefficient:
κ ≤ 0 Less than chance agreement
κ = 0.01-0.20 Slight agreement
κ = 0.21-0.40 Fair agreement
κ = 0.41-0.60 Moderate agreement
κ = 0.61-0.80 Substantial agreement
κ = 0.81-0.99 Almost perfect agreement

Table II. Comparison of inpatient postoperative events in National Surgical Quality Improvement Program (*NSQIP*) and Vascular Quality Initiative (*VQI*) data sets for infrainguinal bypass (total N = 240)

	NSQIP		*VQI*		*κ statistic*	
Variable	*No.*	*%*	*No.*	*%*	*κ coefficient*	*P value*
Mortality						
Yes	3	1.3	3	1.3	1.00	<.0001
No	237	98.8	237	98.8		
MI						
Yes	3	1.3	4	1.7	0.86	<.0001
No	237	98.8	236	98.3		
Pneumonia						
Yes	4	1.7	3	1.3	0.57	<.0001
No	236	98.3	237	98.8		
Change in renal function						
Yes	1	0.4	5	2.1	−0.01	.88
No	239	99.6	235	97.9		
Return to OR						
Yes	8	3.3	6	2.5	0.41	<.0001
No	232	96.7	234	97.5		
Wound infection						
Yes	1	0.4	4	1.7	−0.01	.90
No	239	99.6	236	98.3		

MI, Myocardial infarction; *OR,* operating room.

agreement, whereas race (κ = 0.55) exhibited moderate agreement.

VQI and NSQIP had substantial concordance on the prevalence of diabetes (κ = 0.74), current smoking (κ = 0.69), and hemodialysis (κ = 0.76). There was moderate concordance for ASA classification (κ = 0.59) and presence of COPD (κ = 0.48). Hypertension showed only fair concordance between data sets (κ = 0.35). There was poor concordance for the presence of CHF (κ = 0.14), with VQI recording 32 patients with preoperative CHF and NSQIP recording only four patients. Hypertension, however, had a similar amount of patients diagnosed with the disease in both data sets (204 in NSQIP and 197 in VQI), but the number of identical patients or overlap in each group was low as assessed by the concordance (ie, although the numbers were similar, they did not capture the same patients). Overall, VQI recorded higher rates of current smoking (86% vs 78%) and CHF (13% vs 4%), but NSQIP recorded higher rates of hypertension (85% vs 82%) and diabetes (56% vs 53%). Incidence of hypertension and CHF had the poorest concordance between data sets.

Comparison of postoperative variables between VQI and NSQIP. The mortality and MI rate exhibited almost perfect concordance between data sets (mortality, κ = 1. 00; MI, κ = 0.86) (Table II). Incidence of pulmonary complications and return to the operating room had moderate agreement (κ= 0.57 and 0.41). However, there was poor concordance for wound infection and change in renal function between VQI and NSQIP (κ = −0.01). VQI identified more patients with wound infections (four patients vs one patient) and more patients with changes in renal function (five patients vs one patient) than did NSQIP during the inpatient period.

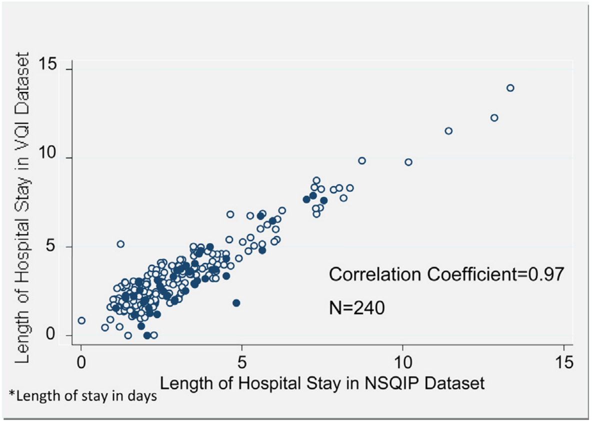

Fig 1. Length of postoperative hospital stay in National Surgical Quality Improvement Program (*NSQIP*) and Vascular Quality Initiative (*VQI*).

Table III. Comparison of inpatient and 30-day postoperative event rates in National Surgical Quality Improvement Program (*NSQIP*) data set (N = 240)

Event	*Inpatient, No. (%)*	*30-day postoperative,[a] No. (%)*
Mortality	3 (75)	1 (25)
MI	3 (50)	3 (50)
Pneumonia	1 (33.3)	2 (66.7)
Return to OR	8 (40)	12 (60)
Change in renal function	1 (50)	1 (50)
Wound infection	1 (4.2)	23 (95.8)
Total	17 (29)	42 (71)

MI, Myocardial infarction; *OR,* operating room.
[a]The 30-day postoperative event does not include any event occurring during the inpatient period.

The median length of stay was 3.0 days (interquartile range, 2-4) in both VQI and NSQIP data sets (Fig 1); Spearman correlation coefficient was 0.97 ($P < .0001$).

Comparison of NSQIP inpatient and 30-day postoperative event rate. Comparison of inpatient events and 30-day postoperative events as reported in the NSQIP data set is shown in Table III. Overall, a total of 59 postoperative events were reported in the NSQIP data set. Whereas 17 (29%) of these events occurred during the inpatient period, 42 (71%) occurred after discharge but within 30 days of the procedure. Specifically, broken down by outcome, three of the four reported deaths (75%) occurred during the inpatient period. Six total occurrences of MI were reported, with an even distribution between the inpatient period (three patients) and the 30-day postoperative period (three patients). Twenty patients required a return visit to the operating room, with a majority (60%) experiencing this during the 30-day

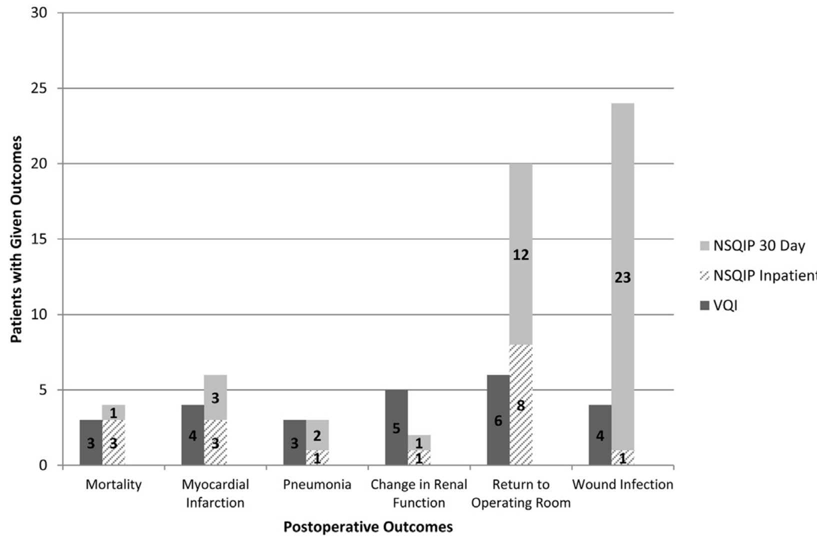

Fig 2. All postoperative events as recorded by National Surgical Quality Improvement Program (*NSQIP*) and Vascular Quality Initiative (*VQI*) in original data set, with stratification for inpatient and 30-day postoperative events in NSQIP.

postoperative period. Only two patients, one in the inpatient setting and the other in the 30-day postoperative period, had a change in renal function. Three cases of pulmonary complication were reported, with only one case occurring during the inpatient period. Last, there were 24 reported cases of wound infection. A significant majority of these cases (95.8%; $P < .0001$) occurred after discharge from the hospital but within 30 days of the surgical procedure. Wound infections also constituted a majority of the 30-day events (54.8%). However, even with the removal of wound infections from the outcomes, 30-day events would still compose a majority (53.8%) of all postoperative events (Fig 2).

DISCUSSION

Our study comparing an identical patient cohort as recorded by two clinical surgical databases, VQI and NSQIP, is the first study of this kind. Despite the use of an identical patient population, the concordance for several preoperative variables and a majority of postoperative variables was fair to poor between these clinical data sets. This finding is significant because the collection of accurate patient information for risk adjustment and outcomes is crucial for benchmarking purposes as well as for calculation of equitable reimbursements and avoidance of inappropriate penalties.

The ACA has set into motion reimbursement guidelines set to take effect shortly. In fiscal year 2013, failure to provide data on specified quality indicators could result in a 2% reduction in Medicare payments to hospitals.[2] The Medicare inpatient prospective payment system will levy additional penalties for risk-adjusted readmission rates greater than the national average, asserting a 1% penalty cap in fiscal year 2013 and a 2% cap in fiscal year 2014. Along the same lines, the ACA hospital-acquired conditions reduction program can also assert a 1% penalty as an incentive to reduce hospital-acquired conditions.[2,3] Under these guidelines, risk adjustment for patient risk factors including comorbidities and outcomes would be used to avoid unfair penalization for hospitals caring for a sicker patient population. These guidelines and penalties are currently instituted for three conditions, MI, CHF, and pneumonia. However, beginning in fiscal year 2015, the Medicare Payment Advisory Commission has recommended the addition of four additional conditions, including vascular conditions.[3] These changes all indicate that accurate data collection of patient demographics and outcomes, to best provide risk-adjusted aggregate data for benchmarking between institutions, is critical.

Comorbidities are important in the decision-making process for patients, yet the concordance for CHF and hypertension was poor and fair, respectively; concordance was the lowest for CHF ($\kappa = 0.13$) among all preoperative variables. This is most likely due to differences in definitions between the two data sets. The guidelines for reporting CHF in NSQIP are stringent and time dependent, requiring a patient to be newly diagnosed with CHF or to have developed new signs and symptoms within the 30 days before intervention (Appendix, online only). VQI, on the other hand, relies on patient history and then further stratifies the level of disease severity, with asymptomatic disease still considered positive for CHF. Not only did this difference result in poor concordance, but VQI reported an eightfold increase in disease burden among an identical patient cohort. This could have significant effects on risk-adjustment methodology and health care policy changes affecting reimbursement. Jencks et al[14] noted that vascular surgery patients had the second highest readmission rate behind CHF patients. Furthermore, as CHF was the third most common cause of those readmissions after a vascular intervention, underreporting and failure to recognize this comorbidity could potentially underestimate the expected number of readmissions and similarly underestimate the expected complication rate. Gupta et al[10] failed to identify CHF as a risk factor for 30-day mortality after infrainguinal bypass using NSQIP data; this may be attributable to the definition used. The CHF rate of 1.9% in Gupta's study was consistent with our findings of 1.7% for the NSQIP cohort, but his finding was much lower compared with that of VQI. McPhee et al[8] specifically reviewed hospital readmission and found that CHF significantly increased the risk of readmission after an infrainguinal bypass procedure. Interestingly, the CHF incidence was 17% in this study, more consistent with our VQI rate of 13.3%. Although clearly assessing different outcomes, these two studies highlight the variability in defining and capturing a key common patient covariate such as CHF.

The discordance in postoperative events may be attributed to both definition and collection methods. Mortality had a perfect concordance, as would be expected given the difficulty in misinterpreting the definition. MI also revealed good concordance, probably attributable to similar definitions between VQI and NSQIP; both databases use a definition that incorporates enzyme levels, electrocardiographic changes, or clinical diagnosis. Pulmonary complications, which included pneumonia and unplanned reintubation, had moderate concordance between data sets. This, like mortality and MI, was probably also due to similar definitions and strict inclusion criteria. Reintubation, like mortality, is a hard end point, leaving little to subjective interpretation. Pneumonia required both a radiologic and a clinical or pathologic finding, creating stricter inclusion criteria, contributing to the moderate concordance observed.

The discordance between postoperative variables becomes more pronounced as the definition relies more heavily on subjective interpretation and the definitions become more dissimilar. Return to the operating room showed a moderate to poor concordance between VQI and NSQIP. Both data sets included only unplanned return to the operating room. In review of the definitions, NSQIP appears to allow more abstractor interpretation than does VQI. According to NSQIP, after answering yes to unplanned trip to the operating room within 30 days, the follow-up question asks if the return trip was "possibly related to the principal operative procedure or concurrent procedure." This leaves the clinical nurse abstractor, although well trained, responsible for making this decision, creating a potential source of interpretation bias. VQI, on the contrary, specifically asks for the reason behind the reintervention, thereby eliminating much of this bias. The true incidence of whether a return trip was planned or unplanned is still left to the discretion of the clinician completing the questionnaire, introducing potential for bias. Wound infections and change in renal function showed complete discordance, $\kappa = -0.01$. Wound infection has a strict definition in NSQIP along with multiple categories (wound disruption, superficial, deep, and organ/space infection), whereas VQI combines all wound infections into a yes or no dichotomous answer. According to NSQIP, several findings would constitute a wound infection, including diagnosis by the surgeon or attending physician and positive wound cultures. However, VQI also includes the addition of an antibiotic regimen in the setting of a possible wound infection as an inclusion criterion; this is not required in NSQIP. Overall, NSQIP allows inclusion based on a series of descriptions, inviting more abstractor interpretation, whereas VQI requires either a positive culture or administration of antibiotics in the setting of physician-diagnosed wound infection. Change in renal function has different numerical values between VQI and NSQIP. NSQIP requires an increase of >2 mg/dL above baseline, whereas VQI denotes change in renal function as an increase of >0.5 above baseline (before 2010, change in renal function was a yes/no answer without preset values). These differences in values accounted for the discordance in change in renal function between NSQIP and VQI.

Discordance between clinical data sets and administrative data has been previously reported, as has comparison of national administrative data sets to multi-institutional registries.[15-17] In reviewing the Nationwide Inpatient Sample, the largest inpatient database publicly available, Hertzer[16] found potential underreporting of preprocedural symptoms as well as variability in coding for complications and comorbidities for patients undergoing carotid endarterectomy or stent procedures compared with other large trials. As previously stated, NSQIP data are collected by surgical clinical reviewers and are routinely audited; inter-rater reliability of NSQIP data shows a disagreement rate of only 1.56% between abstractors.[18] That being said, the correlation between NSQIP data and physician chart review has proved to be less robust. Bensely et al[17] compared NSQIP data with administrative data and found fewer physiologic high-risk patients in the NSQIP data set undergoing carotid intervention than in the physician chart review, which they also attributed to differences in definitions. Despite attempts at validation of each data set as well as comparison to physician chart reviews and administrative data, a direct comparison between NSQIP and VQI has not previously been undertaken. This study provides the first comparison of two data sets on an identical patient population.

The capture of postoperative events has been shown to increase with longer follow-up.[19-21] Slightly more than two thirds of all postoperative events in the NSQIP data set occurred in the 30-day postoperative period after discharge from the inpatient setting. The majority of these complications were wound infections, with a significant proportion (94%) occurring after discharge. This is consistent with previous studies evaluating outcomes in both general surgery and vascular surgery. Sidawy et al[19] reported that 31% of all complications for carotid artery stent procedures occurred within the 30-day period after hospital discharge. This suggests that recording of inpatient complication rates alone fails to capture a significant number of relevant events. Fokkema et al[20] demonstrated similar findings when evaluating carotid endarterectomy patients in the NSQIP data set, with 38% of events occurring after hospital discharge but within 30 days of the initial procedure. These finding are not limited to vascular surgery alone. A recent presentation looking at multiple surgical specialties found that 32% of all complications occurred within 30 days of surgery but after discharge from the inpatient setting.[21] Surgical infections were responsible for more than 50% of all these complications, a finding consistent with our study. Overall, longer follow-up periods, especially up to 30 days, will significantly increase the postoperative event capture rate.

VQI and NSQIP both collect valuable information, including patient comorbidities and perioperative outcomes used to risk stratify and benchmark outcomes, but neither assesses socioeconomic status. The current models being used for perioperative data are comprehensive but remain incomplete. Durham et al[22] found that socioeconomic status played a significant role in advanced presentation, cost per day of patency, and inferior limb salvage, yet most outcomes data sets fail to recognize or record this piece of information. Socioeconomic status remains underreported and may have an equal contribution in determining patients at high risk for postoperative events as well as readmissions. A recent attempt at creating a conceptual model for preventing readmissions highlighted the need to look beyond the standard patient demographic information and included social, economic, and patient education factors, something that is currently lacking in both of these data sets.[23]

There are several limitations inherent to our study design. The event rates for the outcomes we compared were relatively small and therefore easily influenced by minor differences. Also, a formal internal validation was not performed to confirm the results obtained by VQI or NSQIP. Although that was not the premise of the study, it may have provided greater detail to explain the poor concordance seen with some variables. We did not use missing variable imputation methods for this study as we thought it was important to note which values were not recorded in each data set. The rate of missing variables was quite low but nonetheless present and may have affected the concordance for certain variables. Finally, our entire study cohort was composed of patients from a single institution. It is possible that some of the differences related to variable collection may be attributable to institution-specific collection methods or biases, thereby limiting the generalizability of our findings.

CONCLUSIONS

VQI and NSQIP are two valuable data sets that will continue to play important roles in quality improvement. On the basis of their different collection methods, definitions, and data collection time frames, hospitals and surgeons should be aware that the two data sets cannot be used to directly compare risk-adjusted patient outcomes. The ability to consistently collect preoperative variables and surgical outcomes will play an important role in risk stratification, quality improvement, public reporting, and health care reimbursement. Compared with collecting inpatient data alone, a collection period that includes 30-day postoperative follow-up will capture a significantly greater number of clinically relevant events.

AUTHOR CONTRIBUTIONS

Conception and design: FA, AS, AR, LM
Analysis and interpretation: FA, AS, NK
Data collection: FA, BS
Writing the article: FA, BS
Critical revision of the article: FA, AS, AR, WR
Final approval of the article: FA
Statistical analysis: NK, PG
Obtained funding: FA, LM
Overall responsibility: FA

REFERENCES

1. The Patient Protection and Affordability Care Act. 111th United States Congress. March 23, 2010.
2. Hospital acute inpatient service payment. Payment basics. MedPAC. October 2012. P1-P5. Available at: http://www.medpac.gov/documents/MedPAC_Payment_Basics_11_hospital.pdf.
3. Centers for Medicare and Medicaid Services, Department of Health and Human Services. Medicare program; hospital inpatient prospective payment systems for acute care hospitals and the long-term care hospital prospective payment system and Fiscal Year 2014 rates; quality reporting requirements for specific providers; hospital conditions of participation; payment policies related to patient status. Fed Regist 2013;78:50495-51040.
4. Kassin MT, Owen RM, Perez SD, Leeds I, Cox JC, Schnier KS, et al. Risk factors for 30-day hospital readmission among general surgery patients. J Am Coll Surg 2012;215:322-30.
5. Moxey PW, Brownrigg J, Kumar SS, Crate G, Holt PJ, Thompson MM, et al. The BASIL survival prediction model in patients with peripheral arterial disease undergoing revascularization in a university hospital setting and comparison with the FINNVASC and modified PREVENT scores. J Vasc Surg 2013;57:1-7.
6. Schanzer A, Mega J, Meadows J, Samson RH, Bandyk DF, Conte MS. Risk stratification in critical limb ischemia: derivation and validation of a model to predict amputation-free survival using multicenter surgical outcomes data. J Vasc Surg 2008;48:1464-71.
7. Adam DJ, Beard JD, Cleveland T, Bell J, Bradbury AW, Forbes JF, et al. Bypass versus angioplasty in severe ischaemia of the leg (BASIL): multicenter, randomised controlled trial. Lancet 2005;366:1925-34.
8. McPhee JT, Barshes NR, Ho KJ, Madenci A, Ozaki CK, Nguyen LL, et al. Predictive factors for 30-day unplanned readmission after lower extremity bypass. J Vasc Surg 2013;57:955-62.
9. Bertges DJ, Goodney PP, Zhao Y, Schanzer A, Nolan BW, Likosky DS, et al. The Vascular Study Group of New England Cardiac Risk Index (VSG-CRI) predicts cardiac complications more accurately than the Revised Cardiac Risk Index in vascular surgery patients. J Vasc Surg 2010;52:674-83.

10. Gupta PK, Ramanan B, Lynch TG, Sundaram A, MacTaggart JN, Gupta H, et al. Development and validation of a risk calculator for prediction of mortality after infrainguinal bypass surgery. J Vasc Surg 2012;56:372-9.
11. The Vascular Quality Initiative. Available at: http://www.vascularqualityinitiative.org/. Accessed December 2012.
12. American College of Surgeons, National Surgical Quality Improvement Program. Available at: http://site.acsnsqip.org/. Accessed December 2012.
13. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005;37:360-3.
14. Jencks SF, Williams MV, Coleman EA. Rehospitalization among patients in the Medicare fee-for-service program. N Engl J Med 2009;360:1418-28.
15. Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, et al. Risk prediction models for hospital readmission: a systematic review. JAMA 2011;306:1688-98.
16. Hertzer NR. The Nationwide Inpatient Sample may contain inaccurate data for carotid endarterectomy and carotid stenting. J Vasc Surg 2012;55:263-7.
17. Bensley RP, Yoshida S, Lo RC, Fokkema M, Hamdan AD, Wyers MC, et al. Accuracy of administrative data versus clinical data to evaluate carotid endarterectomy and carotid stenting. J Vasc Surg 2013;58:412-9.
18. Shiloach M, Frencher SK, Steeger JE, Rowell KS, Bartzokis K, Tomeh MG, et al. Toward robust information: data quality and inter-rater reliability in the American College of Surgeons National Surgical Quality Improvement Program. J Am Coll Surg 2010;210:6-16.
19. Sidawy AN, Zwolak RM, White RA, Siami FS, Schermerhorn ML, Sicard GA, et al. Risk-adjusted 30-day outcomes of carotid stenting and endarterectomy: results from the SVS Vascular Registry. J Vasc Surg 2009;49:71-9.
20. Fokkema M, Bensley RP, Lo RC, Hamden AD, Wyers MC, Moll FL, et al. In-hospital versus postdischarge adverse events following carotid endarterectomy. J Vasc Surg 2013;57:1568-75.
21. Frangou C. 32% of complications diagnosed after discharge from the hospital, study shows. Findings come amid increased public scrutiny; seeking level playing field in reporting. General Surgery News 2012;39:14-15. Available at: http://www.generalsurgerynews.com/Default.aspx.
22. Durham CA, Mohr MC, Parker FM, Bogey WM, Powell CS, Stoner MC. The impact of socioeconomic factors on outcome and hospital costs associated with femoropopliteal revascularization. J Vasc Surg 2010;52:600-7.
23. Brooke BS, DeMartino RR, Girotti M, Dimick JB, Goodney PP. Developing strategies for predicting and preventing readmissions in vascular surgery. J Vasc Surg 2012;56:556-62.

Submitted Nov 27, 2013; accepted Jan 19, 2014.

Additional material for this article may be found online at www.jvascsurg.org.

INVITED COMMENTARY

Eric D. Endean, MD, *Lexington, Ky*

The authors are to be commended for a timely and unique study that compares the Vascular Quality Initiative (VQI) and National Surgical Quality Improvement Program (NSQIP) data sets in an identical patient cohort. It is widely accepted that risk adjustment is essential when reporting outcomes and assumed that the database used for this assessment accurately reflects patient co-morbidities and outcomes. The results of this study therefore are remarkable in that variables that are seemingly "black and white," such as sex, ethnicity, prevalence of diabetes, or variables that should be extracted from the chart as a "number" (eg, American Society of Anesthesiologists class) showed more variability than might be expected. It is also disturbing that there is a lack of concordance in patient outcomes such as wound infection or change in renal function. Such discordance becomes concerning when patient outcomes are used as a factor for determining reimbursement, financial penalties, and/or rating the practitioner. It is feasible that a provider's outcomes could be in the acceptable range as monitored by one methodology yet be penalized by a payor using another methodology.

The current study compared two databases, but many others are also being used such as data for the University Health Consortium, Society for Thoracic Surgery, disease-specific databases, and specialty-specific databases. Each database has inherent differences that are based on who collects or reports the information; the completeness of the medical record; definitions used for patient conditions, complications, or outcomes; the time frame for collection of data; and its intended purpose such as hospital related information, physician outcome, or disease-specific information. The authors have identified many of these issues as contributing to discordance between datasets. Because of associated costs, the number of programs in which institutions and practices participate will be limited. This naturally leads to questions regarding which database is more accurate or which database should be used to measure performance, specifically as it relates to payment. Studies such as this must be undertaken to address these questions. Physician and specialty society input is essential to guide the development and rational use of database information. It is also imperative that physicians understand the methodology and definitions that are used in any program that measures patient outcomes. The authors are to be commended for this study, and I look forward to future work by this group.

Appendix (online only). Definition of preoperative and postoperative variables as defined in National Surgical Quality Improvement Program (*NSQIP*), Vascular Quality Initiative (*VQI*), and combined analysis

	NSQIP	*VQI*	*Definition used for NSQIP and VQI variables*
Preoperative variables			
Gender	Defined as either male or female	Defined as either male or female	No changes or consolidation made. Yes or no as recorded by each data set.
Ethnicity	Defined as Hispanic or non-Hispanic	Defined as Hispanic or non-Hispanic	No changes or consolidation made. Yes or no as recorded by each data set.
Race	Defined as white, black, American Indian, Pacific Islander, Asian, or unknown. If more than one race listed, first race was used.	Defined as white, black, American Indian, Pacific Islander, Asian, or unknown. If more than one race listed0, first race was used.	Consolidated to a dichotomous variable: white and nonwhite. Nonwhite includes black, American Indian, Pacific Islander, Asian, or unknown.
Diabetes	The treatment regimen of the patient's chronic, long-term management No diabetes: - Categorized as no diabetes if patient did not have diagnosis or was controlled by diet alone Diabetes: - Noninsulin: requiring therapy with a noninsulin antidiabetic agent - Insulin: requiring daily insulin therapy	No diabetes Diabetes: - Diet: controlled with diet regimen alone and not requiring any oral medication or insulin - Noninsulin medications: oral medications and noninsulin injections - Insulin: requires use of injectable insulin	Stratified into three categories NSQIP: No, if no diabetes or controlled with diet; oral if controlled with oral medications, and insulin if controlled with insulin regimen. VQI is categorized as no if no diabetes; oral if controlled with diet or oral medications, and insulin if controlled with insulin.
Smoking	If the patient has smoked cigarettes in the year prior to admission for surgery Patients who smoke cigars or pipes or use chewing tobacco are not included	No smoking Prior: quit smoking >1 year ago Current: still smoking within last 12 months; includes cigarettes, pipe, cigar, smokeless, chewing tobacco	No changes or consolidation made for current smoking. Yes or no as recorded by each data set.
CHF	New diagnosis of CHF or chronic CHF with new onset of symptoms within 30 days of surgery	Categorized as follows: - Asymptomatic - Mild limitation of physical activity - Moderate limitation of physical activity - Severe or inability to carry out any activity without discomfort	Consolidated to a dichotomous variable VQI is yes for all levels of disease burden and no if nonexistent. NSQIP is yes or no as recorded by data set.
COPD	No COPD History of COPD or bronchitis per patient or medical record resulting in one or more of the following: - Functional disability form COPD - Hospitalization for COPD - Requires chronic bronchodilator therapy - FEV_1 <75% on pulmonary function testing - Patients with asthma, diffuse interstitial fibrosis, or sarcoidosis were not included.	No COPD COPD: - Not treated with medication - COPD treated with bronchodilator or steroid medications	Consolidated to a dichotomous variable NSQIP is defined as yes or no as recorded by data set. VQI is yes for all levels of disease burden and no if nonexistent.

(*Continued on next page*)

Appendix (online only). Continued.

	NSQIP	*VQI*	*Definition used for NSQIP and VQI variables*
Hypertension	As documented in patient's medical record and requiring medical treatment within 30 days before operative procedure or at time of consideration for surgery	As documented in history or preoperative blood pressure of >140/90 mm Hg	No changes or consolidation made. Yes or no as recorded by each data set.
Dialysis	Acute or chronic renal failure requiring treatment with peritoneal dialysis, hemodialysis, hemofiltration, hemodiafiltration, or ultrafiltration within 2 weeks before operative procedure	Current treatment with hemodialysis or peritoneal dialysis Patients with history of functioning kidney transplant separately included in this category.	Consolidated to a dichotomous variable Categorized as yes if receiving dialysis in either VQI or NSQIP. Patients with history of functioning kidney transplant as defined by VQI stratified to no-dialysis category.
ASA class	The ASA physical status classification of the patient's present physical condition on a scale of 1 to 5 as it appears on the anesthesia record. Report the most recent assessment. ASA 1 − Normal healthy patient ASA 2 − Patient with mild systemic disease ASA 3 − Patient with severe systemic disease ASA 4 − Patient with severe systemic disease that is a constant threat to life ASA 5 − Moribund patient who is not expected to survive without the operation	The ASA physical status classification of the patient's present physical condition on a scale of 1 to 5 as it appears on the anesthesia record. ASA 1 − Normal healthy patient ASA 2 − Patient with mild systemic disease ASA 3 − Patient with severe systemic disease ASA 4 − Patient with severe systemic disease that is a constant threat to life ASA 5 − Moribund patient who is not expected to survive without the operation	No changes or consolidation made. Results as recorded by NSQIP and VQI data set.
Postoperative variables			
Mortality	Death during operation or postoperative death within 30 days of procedure	Mortality during the inpatient period	Same definition in both VQI and NSQIP.
MI	An acute MI during or within 30 days of surgery, defined as: - Electrocardiographic changes with additional finding of ST elevation, new left bundle branch block, new Q wave in more than two leads - New elevation of troponins greater than three times upper level of reference - Physician diagnosis of MI	No MI Troponin elevation only without creatine kinase MB elevation and without other clinical evidence of MI Electrocardiographic changes or clinical evidence of MI in conjunction with any abnormality of cardiac biomarker consistent with infarction	Consolidated to a dichotomous variable Categorized as yes or no if criteria met in either NSQIP or VQI as defined by variable definition for data set.
Pulmonary complication	Pneumonia within 30 days of surgery as diagnosed on chest radiologic examination and clinical symptoms Unplanned intubation separately defined as placement of endotracheal tube or laryngeal mask airway that was not intended or planned	Pneumonia defined as lobar infiltrate on chest radiograph and growth of recognized pathogen Ventilator defined as need for reintubation and ventilator dependence after initial extubation postoperatively	Consolidated to a dichotomous variable Categorized as yes or no if criteria met in either NSQIP or VQI as defined by variable definition for pneumonia or reintubation.
Change in renal function	Progressive renal insufficiency defined as creatinine increase of >2 mg/dL from preoperative value within 30 days of surgery Acute renal failure defined as need for dialysis within 30 days of surgery	New increase in creatinine of 0.5 mg/dL New dialysis includes peritoneal dialysis, hemodialysis, or hemofiltration.	Consolidated to a dichotomous variable Defined as yes if any of the criteria for progressive renal insufficiency or acute renal failure as defined by NSQIP and VQI.

(*Continued on next page*)

Appendix (online only). Continued.

	NSQIP	*VQI*	*Definition used for NSQIP and VQI variables*
Return to OR	Unplanned return to the OR for a surgical procedure within the 30-day postoperative period	Unplanned return to the OR for a surgical procedure	No changes or consolidation made. Yes or no as recorded by each data set.
Wound infection	Classified as superficial incisional SSI, deep incisional SSI, organ/space SSI, or wound disruption if event occurs within 30 days of procedure	Any wound site infection that was culture positive or required antibiotic treatment	Consolidated to a dichotomous variable Categorized as yes or no if criteria met in either NSQIP or VQI as defined by variable definition for data set.

ASA, American Society of Anesthesiologists; *CHF,* congestive heart failure; *COPD,* chronic obstructive pulmonary disease; *FEV_1,* forced expiratory volume in the first second of expiration; *MI,* myocardial infarction; *OR,* operating room; *SSI,* surgical site infection.